DOCUMENT RESUME

ED 186 929 CS 205 570

AUTHOR White, Edward M.
TITLE The Uneasy Compromise: Bringing Together Testers and Teachers of English.
PUB DATE Mar 80
NOTE 8p.; Paper presented at the Annual Meeting of the Conference on College Composition and Communication (31st, Washington, DC, March 13-15, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cooperation; *English Instruction; Higher Education; *Measurement Techniques; *Testing; Testing Problems; Testing Programs; Writing Instruction

ABSTRACT

Teachers and test professionals must work together for responsible measurement in English to take place. Traditionally, they have refused to do so, a situation that has led to serious problems in at least three areas: research in the teaching of writing, classroom practice in teaching revision of writing, and some testing practices by such national organizations as the College Board, Educational Testing Service, American College Testing, and the National Assessment of Educational Progress. The English Placement Test developed by the faculty in the California State University and College systems represents a successful working model of the uneasy compromise teachers and testers can achieve. (Author/FL)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

Edward M. White
Department of English
California State College
San Bernardino, CA. 92407

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Edward M. White
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED186929

CS205570

# THE UNEASY COMPROMISE:

## BRINGING TOGETHER TESTERS AND TEACHERS OF ENGLISH

By
Edward M. White

Anyone who is responsible for a large-scale English testing program needs to confront the vast gulf that divides teachers from testers. At best, and I will describe some "best" experiences later on, teachers and testers work together uneasily and awkwardly. They must work together for responsible measurement. But, normally they refuse to do so, treasuring their ignorance of each other's work, and retreating into a complacent and narrow professionalism.

The English faculty member knows he or she is basically virtuous and humane. We pride ourselves on the individual attention we give students, on our literary sensitivity and writing skill, on the long and thankless hours we work, and on our efforts to preserve and even contribute to the liberal arts tradition in a crass society. Thus we react with distaste, or horror at the idea that someone should suggest that measurement, or, even worse, multiple-choice testing could apply in our field; we tend to believe that the truly important matters are not only undefinable but unmeasurable. We respond with some comfort to lists, such as the one Stephen Judy includes in The ABC's of Literacy (Oxford, 1980), showing the 14 problems with standardized tests. We do not tend to use his moderate and sensible tone on the topic, however, preferring a combination of self-righteous indignation, misinformation, and complacency. We like to sneer at "multiple-guess" tests and to complain about our unaccountable loss of power to outside testing agencies. This sample from an NCTE affiliate publication has the true English teacher tone about measurement:

> "Testing! Ugh! Why did I say that I would write about testing? I don't want to think about testing. I cringe when I hear that someone wants diagramming to be the test for the ability to expand and reduce sentences. I blanch when confronted with the horror of word relationship items on the PSAT. I flee from questions about

the meaning of test scores and selection indexes. I want to carefully sequester my students and me for at least one semester so that I can watch and guide their learning without the fear of district, or state, or college entrance tests. Talk about the impossible dream!"

Some of us do know about testing, and we tend to focus on its unsavory aspects. We know the dismal history of mental measurement, starting with the eugenics movement; we know about the futile quest for definition and measurement of that will 'o the wisp "general intelligence"; and we know that educational measurement has not in general fulfilled its promise to open up our society's social and racial barriers to advancement through education. But most of us not only remain ignorant of testing and other kinds of measurement but point with considerable pride to our inability to cope with numbers statistics, or charts. In some quarters, it is as scandalous to understand educational research as it would be to associate with a nuclear physicist, or to eat at Burger King, Home of the Whopper.

We also tend to be rather proud of the way we relate to students. We spend vast amounts of time marking their papers, unaware of the unfortunate fact that most students never look past our grades to our comments or abbreviations. Though we often use hostile metaphors to describe this labor (one conscientious colleague of mine spends weekends bleeding all over his student's work), we feel noticeably superior to other members of the faculty because we groan under such a paper load. Some of us do spend considerable time with students in conference, and a fair number of us work with students on prewriting or invention or other matters besides editing. But generally we are absolutely convinced, despite substantial evidence to the contrary, that we produce excellent assignments, and that we are able to evaluate student writing fairly. Finally, most of us have not the faintest idea what test specialists do besides devise stupid multiple-choice tests, but we know they are up to no good and are quite sure they have nothing whatsoever to teach us as we go about our usual business.

Of course, all virtue does not reside with English teachers. The test specialists see themselves as seriously maligned professionals, trying against considerable hostility and prejudice to develop fair and responsible measurement devices. They know they have studied a respectable field, and they have complex, even ambiguous feelings about the relationship of measurement to society. While deep in their hearts they feel that reality only exists if it can be measured, they are keenly aware of the fallibility of tests, and

they develop a statistic called the Standard of Error of Measurement to express that fallibility for each test. When that number is ignored, as it customarily is, they are reminded that the naivete about tests in society in general and among English faculty in particular tends to give test firms more power than they should have or can comfortably wield.

The test specialists are, however, very uneasy about sharing that power with faculty, particularly English faculty. They have data to show how unfair most evaluation of student writing is in the field of English, as elsewhere. They know that multiple-choice testing yi lds much more reliable results than most essay testing, and they have tried in vain numbers of times to talk about both the advantages and disadvantages of these kinds of testing to faculty, who tend to turn them off with a slur or a bad joke. Thus they have lost interest in essay testing, if they ever had any, and prefer the certainty of multiple-choice usage tests. Issues such as whether or not students have the right to their own language strike them as a nuisance and they have little patience for the uncertainties about correctness that seem to bother many English teachers. Reading and writing seem to them quite separable skills, and they do not understand very well why so many English teachers have such passionate attachment to literature. They do work with faculty committees when they must, and try to get the faculty committees to understand what they can most usefully contribute. But generally, they see a faculty committee as a necessary evil, because so few faculty are willing to understand the difficulties of putting together both a reliable and valid test; the faculty generally get in the way when they try to do the measurement easily and efficiently.

Now, it is perfectly clear that if English teaching and English measurement are to improve, this not altogether farfetched set of clichés must somehow be replaced. It is not absurd to imagine a world in which at least some English faculty at all levels have learned something important about measurement, and have come to respect what measurement can add to their teaching and research. Nor is it too much to ask those working for test firms who are in English measurement seriously to respect the faculty and the discipline for what it does bring to students and to tests. The kinds of problems we routinely encounter under the present situation have become intolerable. Let me list three of them.

First, research in the teaching of writing is seriously hampered by the fact that we have so few reliable and valid ways of measuring success or

failure in teaching or in writing. We have heard repeatedly from scholars in the field that problems with the reliability of measurement have hampered or destroyed the credibility of research. Unfortunately, the development of so basic a research tool as accurate measurement has been painfully slow and difficult. We have seen funding organizations such as the Sloan Foundation approach the discipline with ample funds ready for research and for revision of the curriculum only to retreat in dismay at the statistical incompetence and research confusion it encountered. If the field is to progress, we must do better in the area of measurement.

Secondly, we know one of the truisms in the discipline is that rewriting is at least as important as first draft writing. And yet we are woefully inept in getting our students to revise their work. Why? Students cannot improve what they cannot evaluate. Not until we learn how to evaluate writing and help our students learn how to evaluate their own writing can we hope to move past first-draft writing into serious revision.

Thirdly, while many of the more publicized attacks upon testing are not well founded, there are serious problems with national testing firms in relation to the general field of English. These problems can only be dealt with if faculty gain a new kind of foothold at places such as the College Board, the Educational Testing Service (ETS), the American College Testing Program (ACT), and the National Assessment of Educational Progress (NAEP). The following comments about each of these organizations are not intended to be inclusive, but merely to indicate the kinds of issues needing attention:

a) The College Board seems to have grown increasingly remote from faculty. It is more and more in the hands of deans of admissions, counselors, and other non-faculty academicians. To go to a College Board meeting as a faculty member is to be an outsider, rather like a student at the faculty club. The English Disciplinary Committee set up by the Board to review its English testing, turned out (some think to the dismay of some Board officials) to be an active and intelligent committee. When that committee sought to affect SAT policy, the Board was uninterested and unresponsive. The Board continues to offer and promote the Test of Standard Written English (TSWE) despite the opposition of the profession to this disingenuous blending of social class with scholastic aptitude. Despite evidence that the TSWE distributes minorities quite differently than do writing samples, the Board (at least in its Western Regional Office) continues to market that test actively.

2) ETS, which administers much of the testing sponsored by the Board, has been generally the most responsive and responsible testing firm in relation to the English descipline. The College-Level Examination Program (CLEP), despite (or perhaps because of) the bad reputation it gained in the past with its weak General Examination in English Composition, has busily revised this test (and others) using excellent faculty committees. Advanced Placement Program (AP) committees tend to be listened to and respected, and to love their work.

Nonetheless, all is not perfect in Princeton. Even though a respected specialist, Evans Alloway, is directing a new ETS office concerned with writing assessment--his work at various English conferences is only one sign of his committment to the profession--he by no means represents the typical ETS test development person. There is considerable discontent with the new AP English Test, which downplays substantially the role of literature; and AP essay scoring, now reduced to a total of 4 readings for 3 essays, outrages all good practice in relation to essay test reliability as well as a long tradition of responsible practice at ETS. Another example: the National Teachers Examination committee on which I served some years ago was appalled at that test, said so politely, and has received no word of any change since. (I did receive a letter some time after delivering this talk at the CCCC convention, March, 1980, indicating that substantial changes will finally be made in the NTE.)

3) The ACT administers an English Expression test, which many colleges use as an alternative to the College Board SAT verbal. What, we need to ask as a profession, is ACT doing to update that dreary test of English Expression? Where are its faculty committees or its research into problems of writing measurement? Where is there even a hint of essay testing at ACT?

4) As chief reader for the 1979 essay scoring for the NAEP, I need to say that the methods used were not as reliable as they should be. We worked diligently for reader consistency, but NAEP was not willing to pay for two readings for each paper, and hence the reliability of the results are open to questions that the data cannot well answer. Thus I join a series of faculty calling for NAEP to raise its standards of reliability for the data it publishes on national levels of writing ability.

I have spoken here of the need to bring faculty and test professionals together to improve English measurement in three areas: Research in writing, the teaching of writing, and in the performance of national testing firms. Much more could be said in this area, but I hope I have indicated the kinds

of problems the gulf between teachers and testers poses for all of us. I will close with a quick account of the California State University and Colleges English Placement Test (CSUC-EPT), which is a working model of the kind of cooperation that should occur.

Our EPT was developed after a systemwide Task Force had recommended a new kind of writing measure for our students, one that would in fact use the knowledge about linguistics, dialectology, etc., that should inform modern English testing and normally does not. As we began to work, under considerable time and money pressure, we tried quite consciously to bring teachers and testers together in the most constructive way.

Thus, when ETS made the winning (and only) bid and was awarded the contract, it gained the right to pay an English faculty test development committee to put the EPT together. The committee held the upper hand: if it did not approve, and stand behind the test, there would be no test, and all the money ETS had invested in test development would be lost. In addition, we specified the faculty to be appointed, as well as the ETS personnel we would and would not accept.

To its credit, ETS took the risk and invested some $100,000 in a test development effort it could not wholly control. The committee was an excellent and responsible one, and the conflicts that developed were eventually resolved. A new kind of test emerged, including a substantial writing component, perhaps the first in a wholly new generation of tests, and tens of thousands of students have by now well tested that test.

Not all was or is glorious and constructive in the relationship between the faculty committee and ETS in relation to the EPT. But the important fact is that all of us took the hard way out, and it worked. The test itself is an accommodation between various points of view, and the relationship between the committee of faculty and the test specialists represents an uneasy compromise that only occasionally dissolves into the customary clichés. The result has not only been a new test for the CSUC and for ETS, but a new attitude towards testing among English teachers and researchers in our system. Best of all, the EPT is "our" test, not "theirs," and "they have helped us do it well.

No one should hope that the kind of venture I have just sketched will become routine, or can easily be established. Everyone has too much stake in clichés and in a narrow sense of professionalism for sensible cooperation

easily to occur. Nonetheless, the EPT provides a real example of the uneasy compromise of my title, and shows that reasonable professionals can work together to good purpose even in the field of English testing.

This paper was presented at the March 1980 CCCC Convention in Washington, D.C.
Dr. White is Coordinator of the Writing Skills Program for the California State University and Colleges.